

正规化项分类(惩罚项)

例如 : $W = [w_1, w_2, \dots, w_n]$

- L0 : 表示所有非0元素的个数
- L1 (Lasso) : $|W|_1$ 表示所有元素的绝对值和即为 $\sum_{i=1}^n |w_i|$
- L2 (ridge) : $|W|_2$ 表示所有元素的平方和 $\sum_{i=1}^n w_i^2$

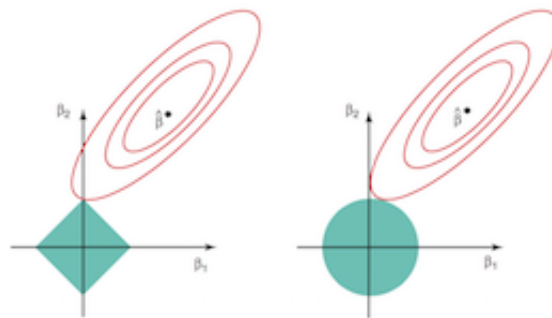
Q : 为什么L1 惩罚项可以将某个元素置0, 而L2不可以?

首先我们公式的形式写出加上了正规化项的目标函数:

$$\text{MIN}_w (y - wx)^2 + L = \text{RSS} + L$$

- 对于**Lasso** : 优化式写为 $\min_{w,b} \{ \sum_{i=1}^n (y_i - wx_1 + b)^2 \}, \sum_{i=1}^n |w_i| \leq s$
- 对于**Ridge** : 优化式写为 $\min_{w,b} \{ \sum_{i=1}^n (y_i - wx_1 + b)^2 \}, \sum_{i=1}^n w_i^2 \leq s$

以二维的向量举例, 最优点表现为圆环(最小二乘法的最优值) 和 绿色阴影的 交点:



左L1(Lasso), 右L2(Ridge)

当某一个惩罚元素置为0的时候, 意味着该最优点必须在坐标轴上, 由于L2图为原形, 椭圆和圆的交点不可能落在坐标轴上, 而合正方形的交点有可能在坐标轴. 所以这就意味着. **L1**的惩罚项元素可能为0, 但是**L2**最多只能无限接近0

因此推广到更高维的空间也是一样.